Assignment 1

Cory J. Cascalheira

01/27/2021

# Assignment Instructions

In a Word document, respond to the following and upload via Canvas:

1. Download the rollercoasters.sav file and open it in SPSS (or import it into R)
2. Write up a response to the following question, using an ANOVA: Do roller coasters that have different excitement ratings differ in their max speed?
   * Make sure to describe center, spread, and shape of the distribution of the variables (Analyze, Descriptives, Explore may be helpful here)
   * Are assumptions of ANOVA met? (What are the assumptions? You can look in your book for this, or also use Laerd Statistics (Links to an external site.))
   * Include all relevant output from the ANOVA in your response to the question

# Establish the Work Environment

Load all the dependencies and import the data.

# Dependencies  
library(car)  
library(psych)  
library(tidyverse)  
library(rio)  
library(lsr)  
library(pander)  
  
# Import  
roller <- import("../data/rollercoasters.sav")  
  
# Construct the APA theme for plots  
# Construct the APA theme  
apa\_theme <- theme\_bw() +  
 theme(panel.grid.major = element\_blank(),  
 panel.grid.minor = element\_blank(),  
 panel.border = element\_blank(),  
 axis.line = element\_line(),  
 plot.title = element\_text(hjust = 0.5),  
 text = element\_text(size = 12, family = "sans"),  
 axis.text.x = element\_text(color = "black"),  
 axis.text.y = element\_text(color = "black"),  
 axis.title = element\_text(face = "bold"))

# ANOVA: Tests of Difference

A helpful R cookbook for executing one-way ANOVAs [can be found here](http://www.sthda.com/english/wiki/one-way-anova-test-in-r). I also briefly consulted Fields et al. (2012). A one-way ANOVA is appropriate because (a) the independent variable (IV) is discrete with three or more groups, (b) the dependent variable (DV) is continuous, and (c) there are no more than two variables, one IV and one DV. It is reasonable to assume that the excitement ratings are independent groups (i.e., between subjects procedure).

It can be helpful to recode the IV as a factor before executing the one-way ANOVA. Note that the effect size is calculated using the type II sum of squares because [the groups are unbalanced](https://www.rdocumentation.org/packages/lsr/versions/0.5/topics/etaSquared).

# Factor recoding  
roller\_1 <- roller %>%  
 mutate(excitement\_rating = as.factor(excitement\_rating))  
  
# Factors in correct order  
roller\_1$excitement\_rating <- ordered(roller\_1$excitement\_rating,   
 levels = c("Low", "Medium", "High", "Very High"))  
  
# Execute ANOVA  
roller\_aov <- aov(max\_speed ~ excitement\_rating, data = roller\_1)  
  
# Print ANOVA results  
summary(roller\_aov)

## Df Sum Sq Mean Sq F value Pr(>F)   
## excitement\_rating 3 2689 896.2 14.24 3.86e-08 \*\*\*  
## Residuals 138 8682 62.9   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Balance of groups  
roller\_1 %>%  
 count(excitement\_rating)

## excitement\_rating n  
## 1 Low 2  
## 2 Medium 12  
## 3 High 116  
## 4 Very High 12

# Calculate effect size  
etaSquared(roller\_aov, type = 2)

## eta.sq eta.sq.part  
## excitement\_rating 0.2364479 0.2364479

The omnibus test is significant, so pairwise comparisons using a Bonferroni-corrected alpha are appropriate. At this point, the calculation of descriptive statistics will be useful for reporting.

# Describe the sample  
roller\_1 %>%  
 group\_by(excitement\_rating) %>%  
 summarize(  
 M = mean(max\_speed),  
 SD = sd(max\_speed),  
 n = n()  
 )

## # A tibble: 4 x 4  
## excitement\_rating M SD n  
## <ord> <dbl> <dbl> <int>  
## 1 Low 31.5 3.54 2  
## 2 Medium 37.9 4.38 12  
## 3 High 43.0 7.25 116  
## 4 Very High 56.4 14.8 12

# Post-hoc pairwise comparisons  
pairwise.t.test(roller\_1$max\_speed, roller\_1$excitement\_rating,  
 p.adjust.methods = "bonferroni")

##   
## Pairwise comparisons using t tests with pooled SD   
##   
## data: roller\_1$max\_speed and roller\_1$excitement\_rating   
##   
## Low Medium High   
## Medium 0.29136 - -   
## High 0.10540 0.10540 -   
## Very High 0.00027 3.9e-07 6.6e-07  
##   
## P value adjustment method: holm

Notably, the *p* adjustment method defaults to Holm, which may indicate a problem with the assumptions (i.e., group size). Nonetheless, without investigating assumptions yet, the results can be reported.

The one-way ANOVA tested the null hypothesis that there was no difference in maximum speed between roller coasters with four different excitement ratings: low, medium, high, and very high. The overall results were significant, *F*(3, 138) = 14.24, *p* < .001, = .24. Pairwise comparisons using the Holm-adjustment method showed three significant differences between very high (*M* = 56.42, *SD* = 14.84) and low (*M* = 31.5, *SD* = 3.54) excitement ratings, *p* < .001; very high and medium (*M* = 37.92, *SD* = 4.38) excitement ratings, *p* < .001; and very high and high (*M* = 43.03, *SD* = 7.25) excitement ratings, *p* < .001. None of the other pairwise comparisons were significant.

## Describe the Distributions

The **center** of the distributions is given by the mean, median, and mode. These values are presented in a table. Since R does not have a built-in function for the mode, [one is provided](https://www.tutorialspoint.com/r/r_mean_median_mode.htm). When these three values are close to each other, there is evidence of normality (Rovai et al., 2014). Unsurprisingly, the mean, median, and mode are closest for the distribution of maximum speed scores for the High excitement rating, which is reasonable since the number of observations is greatest for this subset of scores (*n* = 116).

# Funciton for mode  
# Create the function.  
getmode <- function(v) {  
 uniqv <- unique(v)  
 uniqv[which.max(tabulate(match(v, uniqv)))]  
}  
  
# Calculate measures of center  
roller\_1 %>%  
 group\_by(excitement\_rating) %>%  
 summarize(  
 Mean = mean(max\_speed),  
 Median = median(max\_speed),  
 Mode = getmode(max\_speed)  
 ) %>%  
 rename(Rating = excitement\_rating) %>%  
 pandoc.table()

##   
## -----------------------------------  
## Rating Mean Median Mode   
## ----------- ------- -------- ------  
## Low 31.5 31.5 34   
##   
## Medium 37.92 39 39   
##   
## High 43.03 42.5 41   
##   
## Very High 56.42 51.5 53   
## -----------------------------------

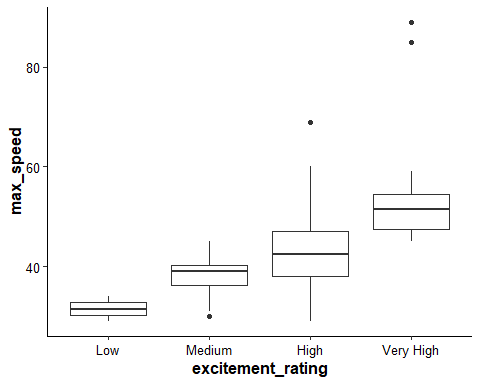
The **spread** (i.e., dispersion) of the distributions can be given by the range, interquartile range (IQR), variance, and standard deviation. These values are presented in a table, demonstrating that the spread increases along with the ratings. That is, the smallest amount of spread exists in the distribution of maximum speed for Low excitement rating while the largest amount of spread is in the distribution of maximum speed for Very High excitement rating.

# Table of dispersion  
roller\_1 %>%  
 group\_by(excitement\_rating) %>%  
 summarize(  
 Max = max(max\_speed),  
 Min = min(max\_speed),  
 Range = Max - Min,  
 IQR = IQR(max\_speed),  
 Variance = var(max\_speed),  
 SD = sd(max\_speed)  
 ) %>%  
 rename(Rating = excitement\_rating) %>%  
 pandoc.table()

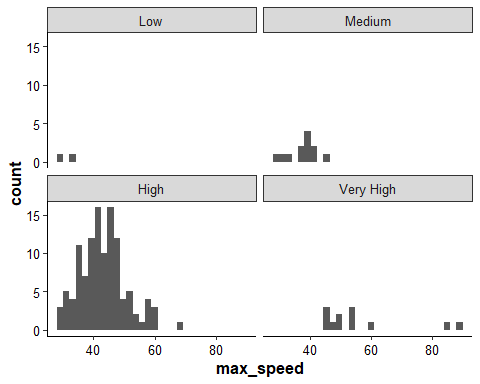
##   
## --------------------------------------------------------  
## Rating Max Min Range IQR Variance SD   
## ----------- ----- ----- ------- ----- ---------- -------  
## Low 34 29 5 2.5 12.5 3.536   
##   
## Medium 45 30 15 4 19.17 4.379   
##   
## High 69 29 40 9 52.5 7.246   
##   
## Very High 89 45 44 7 220.1 14.84   
## --------------------------------------------------------

The **shape** of the distributions can be visualized with boxplots, histograms, and density curves.

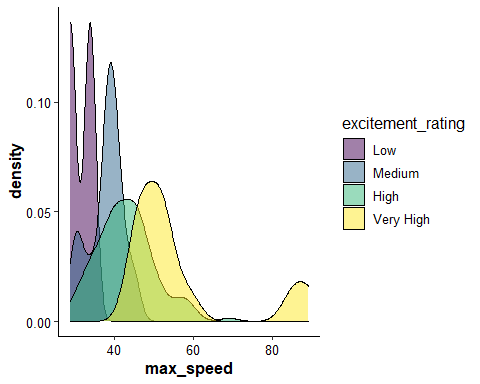
# Show a boxplot  
roller\_1 %>%  
 ggplot(aes(x = excitement\_rating, y = max\_speed)) +  
 geom\_boxplot() +  
 apa\_theme



# Show a histogram of each distribution  
roller\_1 %>%  
 ggplot(aes(x = max\_speed)) +  
 geom\_histogram() +  
 facet\_wrap(~ excitement\_rating) +  
 apa\_theme



# Show the density curve of each distribution  
roller\_1 %>%  
 ggplot(aes(x = max\_speed, fill = excitement\_rating)) +  
 geom\_density(alpha = 0.5) +  
 apa\_theme



Based on a visual inspection of the boxplot, the Medium, High, and Very High excitement ratings have outliers, which will skew the distributions. A visual inspection provided additional insights on the distribution of maximum speed:

1. For the Low excitement rating, the distribution appears bimodal, symmetric, and not normal, which is expected given that the distribution is comprised of two values.
2. For the Medium excitement rating, the distribution approaches a bimodal shape due to the outliers. Based on the histogram alone, the distribution could be considered unimoda and asymmetric with a negative skew. According to the density plot, the distribution seems leptokurtic.
3. For the High excitement rating, the distribution approaches a unimodal, symmetric, bell shape, although there is evidence of positive skew. Visually, the distribution seems mesokurtic.
4. For the Very High excitement rating, the distribution might be considered unimodal and asymmetric, although this is unclear because the outliers seem to produce a second modal feature in the density plot. There is a clear positive skew. Kurtosis values, once calculated, might support heavy-tailedness (Rovai et al., 2014).

# ANOVA: Assumptions

Whether these results are statistically valid partially depends on the tenability of the assumptions (Rovai et al., 2014). A description of the center, spread, and shape of the distributions revealed several concerns.

The assumptions of a between-subjects, one-way ANOVA are (Fields et al., 2012; Rovai et al., 2014):

1. characteristics of variables;
2. independence;
3. group-level univariate normality;
4. homogeneity of variance; and
5. sample size.

## Characteristics of Variables

The variables meet this assumption. The IV is discrete with more than three levels. The DV is continuous, existing at least at the interval level because the interval between 36 mph and 37 mph is the same as the interval between 88 mph and 89 mph. The distance between values make sense numerically. The DV is probably at the ratio level of measurement because the maximum speed of a broken roller coaster would be zero. Although reported as an integer in this dataset, the maximum speed also likely shifts slightly for each ride.

## Independence

Although we would need more information to decide on this assumption definitively, it is reasonable to suggest that the excitement ratings were made by different people.

Of course, online ratings can have a crowd effect (i.e., a person may rate similar to others for social desirability if ratings are posted). However, for the sake of this exercise, I assumed that the roller coasters were independently rated.

## Group-Level Univariate Normality

There are several ways to assess univariate normality. The assumption is tenable when a package of evidence suggests normality.

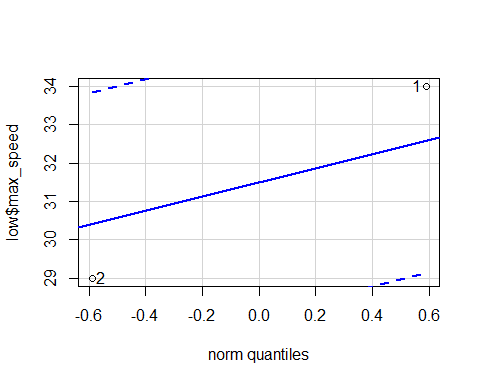
# Vector of values  
ratings <- as.character(unique(roller\_1$excitement\_rating))  
  
# Describe the data  
for (i in 1:length(ratings)) {  
   
 # Select the rating  
 a <- roller\_1 %>%  
 filter(excitement\_rating == ratings[i])  
   
 # Describe the maximum speed  
 b <- describe(a$max\_speed)  
 print(b)  
}

## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 116 43.03 7.25 42.5 42.63 6.67 29 69 40 0.6 0.62 0.67  
## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 12 37.92 4.38 39 38 2.97 30 45 15 -0.42 -0.88 1.26  
## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 12 56.42 14.84 51.5 54.3 6.67 45 89 44 1.35 0.17 4.28  
## vars n mean sd median trimmed mad min max range skew kurtosis se  
## X1 1 2 31.5 3.54 31.5 31.5 3.71 29 34 5 0 -2.75 2.5

The skewness is close to zero for the Medium and Low excitement ratings, but closer to 1 for the High and Very High excitement ratings, suggesting positive skew. These skew values are acceptable, especially for a robust proecedure like an ANOVA (Rovai et al., 2014). The kurtosis value for the Very High excitement rating is close to zero, which matches the mesokurtic observation made earlier. The kurtosis for the High excitement rating is closer to 1, suggesting a leptokurtic distribution; for the Medium rating, the kurtosis is close to -1, suggesting a platykurtic distribution. Nonetheless, these kurtosis values are acceptable. The final kurtosis value suggests non-normality because it is less than -2 (Rovai et al., 2014).

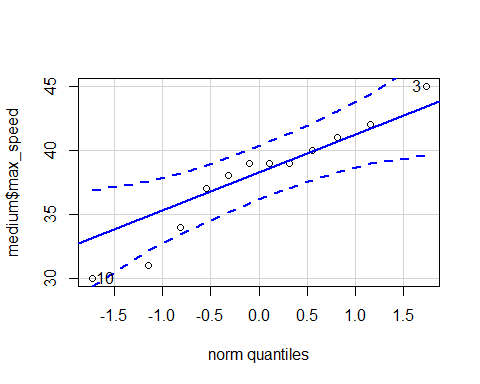
Q-Q plots can be examined.

# Subset the data frame  
low <- roller\_1 %>%  
 filter(excitement\_rating == "Low")  
  
medium <- roller\_1 %>%  
 filter(excitement\_rating == "Medium")  
  
high <- roller\_1 %>%  
 filter(excitement\_rating == "High")  
  
vh <- roller\_1 %>%  
 filter(excitement\_rating == "Very High")  
  
# Q-Q plots  
qqPlot(low$max\_speed)



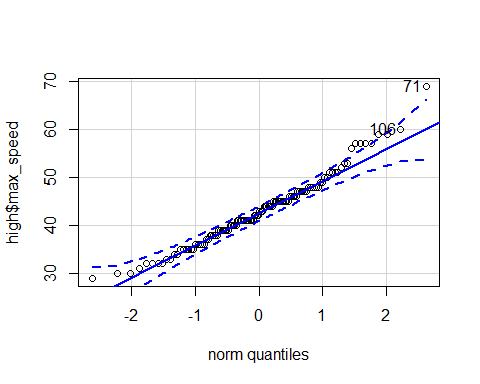
## [1] 2 1

qqPlot(medium$max\_speed)



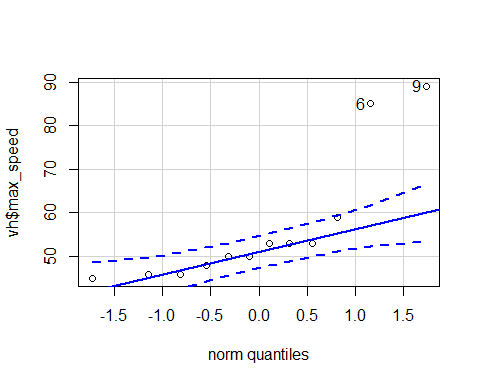
## [1] 10 3

qqPlot(high$max\_speed)



## [1] 71 106

qqPlot(vh$max\_speed)



## [1] 9 6

A visual inspection of the Q-Q plots indicates non-normality for the distribution of maximum speed scores for the Low, Medium, and Very High excitement ratings because the data are asymmetrical along the diagonal line. There is also some asymmetry in the Q-Q plot of maximum speed scores for the High excitement rating, namely, in the upper-right quadrant of the graph.

Formal tests of normality can be conducted as well. The Shapiro-Wilk test is appropriate for *n* < 50 (i.e., Low, Medium, Very High). When *n* > 50, the Kolmogorov-Smirnov (KS) test is preferred (i.e., High). Note that the Shapiro-Wilk test cannot be conducted on a sample of *n* < 3, so it will not be conducted for the Low excitement rating.

# Remove the Low excitement rating  
roller\_2 <- roller %>%  
 filter(!(excitement\_rating %in% c("Low", "High")))  
  
# Shapiro-Wilk tests  
by(roller\_2$max\_speed, roller\_2$excitement\_rating, shapiro.test)

## roller\_2$excitement\_rating: Medium  
##   
## Shapiro-Wilk normality test  
##   
## data: dd[x, ]  
## W = 0.94001, p-value = 0.4983  
##   
## ------------------------------------------------------------   
## roller\_2$excitement\_rating: Very High  
##   
## Shapiro-Wilk normality test  
##   
## data: dd[x, ]  
## W = 0.70352, p-value = 0.0009112

# KS test  
ks.test(x = vh$max\_speed, y = "pnorm")

##   
## One-sample Kolmogorov-Smirnov test  
##   
## data: vh$max\_speed  
## D = 1, p-value = 7.55e-11  
## alternative hypothesis: two-sided

Results from the Shapiro-Wilk test yielded evidence for this distribution of maximum speed scores for the Medium rating, *W* = 0.94, *p* = .50, but not the Very High rating, *W* = 0.70, *p* < .001. The KS test yielded evidence for a non-normal distribution of maximum speed for the Very High rating, *D*(12) = 1, *p* < .001.

In summary, the distribution of maximum speed scores:

1. for the Low excitement rating is probably not normal;
2. for the Medium excitement rating is somewhat normal, but probably not normal enough given the small sample;
3. for the High excitement rating is probably normal; and
4. for the Very High excitement rating is probably not normal.

## Homogeneity of Variance

Levene’s test provides evidence for the equality of variances. Since the distribution of maximum speed for three of the excitement ratings is probably not normal, Levene’s test will be executed using [the median will be used](https://www.itl.nist.gov/div898/handbook/eda/section3/eda35a.htm).

# Levene's test  
leveneTest(roller\_1$max\_speed,   
 group = roller\_1$excitement\_rating,   
 center = median)

## Levene's Test for Homogeneity of Variance (center = median)  
## Df F value Pr(>F)   
## group 3 2.4556 0.06571 .  
## 138   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Compare to centering on the mean  
leveneTest(roller\_1$max\_speed,   
 group = roller\_1$excitement\_rating,   
 center = mean)

## Levene's Test for Homogeneity of Variance (center = mean)  
## Df F value Pr(>F)   
## group 3 5.0048 0.00252 \*\*  
## 138   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The variances among the different excitement ratings were similar when centered on the median, *F*(3, 138) = 2.46, *p* = .07, but significantly different when centered on the mean, *F*(3, 138) = 5, *p* = .002. Given that the sample sizes are quite diverse, it does not seem to make sense that the variances would be equal.

## Sample Size

The sample sizes are not equal across groups, which may increase the Type I error rate (Rovai et al., 2014). Thus, it was likely that the one-way ANOVA returned false-positive significance.

# Evidence of unequal sample sizes  
roller\_1 %>%  
 count(excitement\_rating)

## excitement\_rating n  
## 1 Low 2  
## 2 Medium 12  
## 3 High 116  
## 4 Very High 12

In summary, there is evidence of assumption violation. I would probably not trust the results of this one-way ANOVA because (a) some of group-level distributions are not normal, (b) the sample sizes are quite variable, and (c) there is evidence of unequal variances when the distributions are centered on the mean.

# References

Field, A., J. Miles, & Z. Field. (2012). *Discovering statistics using R*. SAGE Publications Ltd.

Rovai, A. P., J. D. Baker, & M. K. Ponton. (2014). *Social science research design and statistics: A practitioner’s guide to research methods and IBM SPSS analysis* (2nd ed). Watertree Press.